

## AMENDMENTS TO THE CLAIMS

The following list of claims replaces all prior versions and lists of claims:

1. (Currently Amended) A method, comprising:
  - receiving a request on a DNS server from a user client for a web page at a first web address, the first web address including a hostname;
  - determining traffic loads of a plurality of mirrored customer web servers each addressable by the requested hostname among a customer's plurality of web servers, each of the customer web servers storing the web page;
  - determining a customer web server from the plurality of mirrored customer web servers that is appropriate for the request, the customer web server having a traffic load lower than traffic loads of remaining customer web servers from the plurality of mirrored customer web servers;
  - determining an IP address of the customer web server;
  - ~~directing the request from~~ sending the IP address of the customer web server to the client user to the customer web server;
  - receiving a request from the user client for static content on the web page at a second web address, the second web address specifying a network of caching servers including the hostname;
  - determining service metrics of a set of caching servers each addressable by the second web address in [[a] the network of caching servers different from , the network of caching servers does not include the customer's plurality of web servers;
  - wherein a customer ~~pays a fee to~~ is a customer of a service for use of the network of caching servers managed by the service that store ~~storing~~ static content for the customer;
  - determining a caching server from the ~~network set~~ set of caching servers that is appropriate for the request for static content, the caching server having service metrics better than service metrics of remaining caching servers from the ~~network set~~ set of caching servers;
  - determining an IP address of the caching server; and
  - delivering the IP address of the caching server to the client.

~~retrieving the static content from the caching server; and  
providing the static content to the user.~~

2. (Previously Presented) The method of claim 1 further comprising:  
determining load of caching servers in the network of caching servers;  
wherein the determining the caching server from the network of caching servers that is appropriate for the request step selects a caching server having a latency and a load lower than latency or load of remaining caching servers from the network of caching servers.
  
3. (Currently Amended) The method of claim 1 further comprising:  
determining whether the caching server includes the static content;  
determining a customer web server that includes the static content when the caching server does not include the static content;  
retrieving the static content from the customer web server that includes the static content; and  
storing the static content from the customer web server in the caching server.
  
4. (Currently Amended) The method of claim 3 wherein the determining the customer web server that includes the static content step comprises:  
determining traffic loads of the plurality of mirrored customer web servers, each of the customer web servers storing the static content; and  
determining a second customer web server from the plurality of mirrored customer web servers that is appropriate for the request, the second customer web server having a traffic load lower than traffic loads of remaining customer web servers from the plurality of mirrored customer web servers.

5. (Currently Amended) The method of claim 4 wherein retrieving the static content from the customer web server step comprises:

determining an IP address of the second customer web server; and  
requesting the static content from the second customer web server at the second customer web server IP address.

6. (Previously Presented) The method of claim 1 wherein the network of caching servers includes a domain name server.

7. (Currently Amended) The method of claim 1 wherein the request from the ~~user~~ client for the web page is transferred from a first domain name server;

wherein the network of caching servers includes a second domain name server; and

wherein the second domain name server determines the customer web server from the plurality of mirrored customer web servers.

8. (Currently Amended) A method, comprising:  
receiving a first request on a DNS server from a client DNS server to resolve a first domain name, the client DNS server receiving a request from a ~~user~~ client of a web page address that includes the first domain name;

determining load measurements of a plurality of mirrored customer web servers each addressable by the first domain name among a customer's plurality of web servers, each of the customer web servers addressable by the first domain name, and each of the customer web servers configured to service the request from the ~~user~~ client;

determining a customer web server from the plurality of mirrored customer web servers, the customer web server having a traffic load lower than traffic loads of other customer web servers from the plurality of mirrored customer web servers;

determining an IP address of the customer web server;

providing the IP address of the customer web server to the client DNS server;

receiving a second request from the client DNS server to resolve a second domain name, the client DNS server receiving a request from the ~~user~~ client of a uniform resource locator obtained from the web page associated with the web page address that includes the second domain name;

determining performance metric measurement of a set of caching servers each addressable by the second domain name in a network of caching servers ~~different from~~ , the network of caching servers does not include the customer's plurality of web servers, ~~each of the caching servers addressable by the second domain name~~;

wherein a customer ~~pays a fee to~~ is a customer of a service for use of the network of caching servers managed by the service that store ~~storing~~ static content for the customer;

determining a caching server from the ~~network~~ set of caching servers, the caching server having performance metrics lower than performance metrics of other caching servers from the ~~network~~ set of caching servers; and

~~providing~~ delivering an IP address of the caching server to the client DNS server[[]].

~~retrieving data from the caching server in response to the uniform resource locator; and~~

~~providing the data to the user.~~

9. (Original) The method of claim 8 wherein the load measurements comprise latency measurements.

10. (Previously Presented) The method of claim 8 wherein the performance metric measurements comprise any of: load CPU and memory measurements, HTTP response measurements, and FTP response measurements.

11. (Currently Amended) The method of claim 8 further comprising ~~wherein retrieving data from the caching server step comprises:~~

in response to receiving the uniform resource locator request at the caching server, determining whether the caching server includes the data;

retrieving data from a second customer web server from the mirrored customer web servers when the caching server does not include the data; and

storing the data within the caching server.

12. (Previously Presented) The method of claim 11 wherein retrieving data from the second customer web server step comprises:

determining the second customer web server from the plurality of mirrored customer web servers, the second customer web server having a traffic load lower than traffic loads of remaining customer web servers from the plurality of mirrored customer web servers; and

retrieving the data from the second customer web server.

13. (Currently Amended) The method of claim 8 further comprising: receiving a first request from a second client DNS server to resolve a third domain name, the second client DNS server receiving a request from a second ~~user~~ client of a second web page address that includes the third domain name;

determining load measurements of a plurality of second mirrored customer web servers each addressable by the third domain name among a customer's plurality of web servers, ~~each of the second customer web servers addressable by the third domain name,~~ and each of the second customer web servers storing data configured to service the request from the second ~~user~~ client;

determining a second customer web server from the plurality of second mirrored customer web servers, the second customer web server having a traffic load lower than traffic loads of other second customer web servers from the plurality of second mirrored customer web servers;

determining an IP address of the second customer web server; and

providing the IP address of the second customer web server to the second client DNS server.

14. (Currently Amended) The method of claim 13 further comprising: receiving a second request from the second client DNS server to resolve the second domain name, the second client DNS server receiving a request from the second ~~user~~ client of a second uniform resource locator that includes the second domain name;

determining performance metric measurement of the set of caching servers;

determining a second caching server from the set of caching servers, the second caching server having performance metrics lower than performance metrics of other caching servers from the set of caching servers; and

delivering an IP address of the second caching server to the second client DNS server.

~~retrieving a second set of data from the caching server in response to the second uniform resource locator; and~~

~~providing the second set of data to the user.~~

15-20. (Canceled)

21. (Currently Amended) An apparatus, comprising:

~~a module for receiving~~ a DNS server that receives a request from a ~~user~~ client for a web page at a first web address, the first web address including a hostname;

~~a module for determining~~ a traffic load subsystem that determines traffic loads of a plurality of mirrored customer web servers each addressable by the requested hostname among a customer's plurality of web servers, each of the customer web servers storing the web page;

~~a module for determining~~ a customer web server selection subsystem that determines a customer web server from the plurality of mirrored customer web servers

that is appropriate for the request, the customer web server having a traffic load lower than traffic loads of remaining customer web servers from the plurality of mirrored customer web servers;

~~a module for determining~~ an IP address subsystem that determines an IP address of the customer web server;

~~a module for directing~~ a request directing subsystem that directs the request from the ~~user~~ client to the customer web server;

~~a module for receiving~~ a request receiving subsystem that receives a request from the ~~user~~ client for static content on the web page at a second web address, the second web address including the hostname;

~~a module for determining~~ a service metric subsystem that determines service metrics of a set of caching servers each addressable by the requested hostname in a network of caching servers different from , the network of caching servers does not include the customer's plurality of web servers;

wherein a customer ~~pays a fee to~~ is a customer of a service for use of the network of caching servers managed by the service that store ~~storing~~ static content for the customer;

~~a module for determining~~ a caching server selection subsystem that determines a caching server from the network set of caching servers that is appropriate for the request for static content, the caching server having service metrics better than service metrics of remaining caching servers from the network of caching servers;

a caching server IP address determining subsystem that determines an IP address of the caching server; and

a caching server IP address delivery subsystem that delivers the IP address of the caching server to the client.

~~a module for retrieving the static content from the caching server; and~~

~~a module for providing the static content to the user.~~

22. (Currently Amended) The apparatus of claim 21 further comprising:

~~a module for determining~~ a caching server load determination subsystem that determines load of caching servers in the network set of caching servers;

wherein ~~determining the caching server from the network of caching servers that is appropriate for the request module~~ the caching server load determination subsystem selects a caching server having a latency and a load lower than latency or load of remaining caching servers from the ~~network set~~ of caching servers.

23. (Currently Amended) The apparatus of claim 21 further comprising:

~~a module for determining~~ a caching server content determination subsystem that determines whether the caching server includes the static content;

~~a module for determining~~ a customer web server content determination subsystem that determines a customer web server that includes the static content when the caching server does not include the static content;

~~a module for retrieving~~ a customer web server content retrieval subsystem that retrieves the static content from the customer web server that includes the static content; and

~~a module for storing~~ a caching server content storage subsystem that stores the static content from the customer web server in the caching server.

24. (Currently Amended) The apparatus of claim 23 wherein the ~~determining the web server module~~ customer web server content determination subsystem comprises:

~~a module for determining~~ a customer web server traffic load determination subsystem that determines traffic loads of the plurality of mirrored customer web servers, each of the customer web servers storing the static content; and

~~a module for determining~~ a secondary customer web server selection subsystem that determines a second customer web server from the plurality of mirrored customer web servers that is appropriate for the request, the second customer web server having a traffic load lower than traffic loads of remaining customer web servers from the plurality of mirrored customer web servers.



25. (Currently Amended) The apparatus of claim 24 wherein the ~~retrieving the static content from the web server module~~ customer web server content retrieval subsystem comprises:

~~a module for determining~~ a secondary customer web server IP address determination subsystem that determines an IP address of the second customer web server; and

~~a module for requesting~~ a secondary server content requesting subsystem that requests the static content from the second customer web server at the second customer web server IP address.

26. (Previously Presented) The apparatus of claim 21 wherein the network of caching servers includes a domain name server.

27. (Currently Amended) The apparatus of claim 21

wherein the ~~user~~ client for the web page is transferred from a first domain name server;

wherein the network of caching servers includes a second domain name server; and

wherein the second domain name server determines the customer web server from the plurality of mirrored customer web servers.

28. (Currently Amended) An apparatus, comprising:

~~a module for receiving~~ a DNS server that receives a first request from a client DNS server to resolve a first domain name, the client DNS server receiving a request from a ~~user~~ client of a web page address that includes the first domain name;

~~a module for determining~~ a traffic load subsystem that determines load measurements of a plurality of mirrored customer web servers each addressable by the first domain name among a customer's plurality of web servers, ~~each of the customer web~~

~~servers addressable by the first domain name~~, and each of the customer web servers configured to service the request from the ~~user~~ client;

~~a module for determining~~ a customer web server selection subsystem that determines a customer web server from the plurality of mirrored customer web servers, the customer web server having a traffic load lower than traffic loads of other customer web servers from the plurality of mirrored customer web servers;

~~a module for determining~~ an IP address subsystem that determines an IP address of the customer web server;

~~a module for providing~~ wherein the IP address subsystem sends the IP address of the customer web server to the client DNS server;

~~a module for receiving~~ wherein the DNS server receives a second request from the client DNS server to resolve a second domain name, the client DNS server receiving a request from the ~~user~~ client of a uniform resource locator obtained from the web page associated with the web page address that includes the second domain name;

~~a module for determining~~ a caching server performance metric subsystem that determines performance metric measurement of a set of caching servers each addressable by the second domain name in a network of caching servers ~~different from~~ the network of caching servers does not include the customer's plurality of web servers; ~~each of the caching servers addressable by the second domain name~~;

wherein a customer ~~pays a fee to~~ is a customer of a service for use of the network of caching servers managed by the service that store ~~storing~~ static content for the customer;

~~a module for determining~~ a caching server selection subsystem that determines a caching server from the ~~network~~ set of caching servers, the caching server having performance metrics lower than performance metrics of other caching servers from the ~~network~~ set of caching servers; and

~~a module for providing~~ a caching server selection subsystem that delivers an IP address of the caching server to the client DNS server~~[[;]]~~.

~~a module for retrieving data from the caching server in response to the uniform resource locator~~; and

~~a module for providing the data to the user~~.

29. (Previously Presented) The apparatus of claim 28 wherein the load measurements comprise latency measurements.

30. (Previously Presented) The apparatus of claim 28 wherein the performance metric measurements comprise any of: load CPU and memory measurements, HTTP response measurements, and FTP response measurements.

31. (Currently Amended) The apparatus of claim 28 further comprising ~~wherein retrieving data from the caching server comprises:~~

~~a module for determining~~ a caching server data determination subsystem that, in response to receiving the uniform resource locator request at the caching server, determines whether the caching server includes the data;

~~a module for retrieving~~ a second customer web server data retrieval subsystem that retrieves data from a second customer web server from the plurality of mirrored customer web servers when the caching server does not include the data; and

~~a module for storing~~ a caching server data storing subsystem that stores the data within the caching server.

32. (Currently Amended) The apparatus of claim 31 wherein the ~~retrieving data from the second customer web server module~~ second customer web server data retrieval subsystem comprises:

~~a module for determining~~ a second customer web server determining subsystem that determines the second customer web server from the plurality of mirrored customer web servers, the second customer web server having a traffic load lower than traffic loads of remaining customer web servers from the plurality of mirrored customer web servers; and

~~a module for retrieving~~ a second customer web server data retrieval subsystem that retrieves the data from the second customer web server.

33. (Currently Amended) The apparatus of claim 28 ~~further~~ comprising:

~~a module for receiving~~ wherein the DNS server receives a first request from a second client DNS server to resolve a third domain name, the second client DNS server receiving a request from a second ~~user~~ client of a second web page address that includes the third domain name;

~~a module for determining~~ wherein the traffic load subsystem determines load measurements of a plurality of second mirrored customer web servers each addressable by the third domain name among a customer's plurality of web servers, ~~each of the second customer web servers addressable by the third domain name,~~ and each of the second customer web servers storing data configured to service the request from the second ~~user~~ client;

~~a module for determining~~ wherein the customer web server selection subsystem determines a second customer web server from the plurality of second mirrored customer web servers, the second customer web server having a traffic load lower than traffic loads of other second customer web servers from the plurality of second mirrored customer web servers;

~~a module for determining~~ wherein the IP address subsystem detemines an IP address of the second customer web server; and

~~a module for providing~~ wherein the IP address subsystem sends the IP address of the second customer web server to the second client DNS server.

34. (Currently Amended) The apparatus of claim 33 ~~further~~ comprising:

~~a module for receiving~~ wherein the DNS server receives a second request from the second client DNS server to resolve the second domain name, the second client DNS server receiving a request from the second ~~user~~ client of a second uniform resource locator that includes the second domain name;

wherein the caching server performance metric subsystem determines performance metric measurement of the set of caching servers;

wherein the caching server selection subsystem determines a second caching server from the set of caching servers, the second caching server having performance metrics lower than performance metrics of other caching servers from the set of caching servers; and

wherein the caching server selection subsystem delivers an IP address of the second caching server to the second client DNS server.

~~a module for retrieving a second set of data from the caching server in response to the second uniform resource locator; and~~

~~a module for providing the second set of data to the user.~~